A comparison between VGG16, VGG19 and ResNet50 architecture frameworks for Image Classification

Year: 2021 | Citations: 377 | Authors: Sheldon Mascarenhas, Mukul l Agarwal

---

Abstract

Artificial Intelligence advancements have come a long way over the past twenty years. Rapid developments in AI have given birth to a trending topic called machine learning. Machine learning enables us to use algorithms and programming techniques to extract, understand and train data. Machine learning led to the creation of a concept called deep learning which uses algorithms to create an artificial neural network and use it to develop and learn, based on which it makes intuitive decisions by itself. Image classification is a task where we classify the images into sets of different categories, which when performed using deep learning increases business productivity by saving time and manpower. In this paper, we intend to determine which model of the architecture of the Convoluted Neural Network (CNN) can be used to solve a real-life problem of product classification to help optimize pricing comparison. We have compared the VGG16, VGG19, and ResNet50 architectures based on their accuracy while all three of these models solve the same image classification problem. We have concluded that the ResNet50 is the best architecture based on the comparison. These models have provided accuracies of 0.9667, 0.9707, and 0.9733 for VGG16, VGG19, and ResNet50 at epoch 20. The data provided is a real-life data set, sourced from a regional retailer.